



PEX Rule Editor

The KantanMT PEX Rule Editor enables members of KantanMT.com to create, test and deploy PEX (Post-Editing Automation) rules on specific client profiles. The editor allows users to evaluate the output of a PEX rule on a sample of translated content without needing to upload it to a client profile and run translation jobs. Users can enter up to three pairs of search and replace rules, which will be run in descending order on your content

Automatic post-editing

As increasing numbers of Language Service Providers (LSPs) and Enterprises integrate Machine Translation into their workflows, a larger emphasis is being placed on the post-editing process.

Post editing is the process of improving Machine Translated output to meet specific quality standards agreed between the client and the LSP, or language department. The post-editing process can involve three steps; *automated post editing, human post editing and proof reading.*

KantanMT's PEX Rule Editor

The PEX editor makes the automated post-editing step as easy and seamless as possible.

Using the PEX editor KantanMT members can:

1. Design, develop, and debug post-editing rules
2. Test post-editing rules using sample content
3. Apply these rules to client translation files



PEX Rule Editor

PEX Rule Editor

The PEX Rule Editor uses 'find and replace' rules to automatically fix repetitive errors within a document.

Examples of repetitive errors

Some users of Machine Translation may experience the following errors in their MT output.

- Capitalisation
- Punctuation mistakes, hyphenation, diacritic marks etc.
- Words added/omitted
- Formatting – trailing spaces

The introduction of automated post-editing tools significantly speeds up the correction of these common errors, and helps to relieve post-editors of some of the more repetitive post-editing tasks.

Regular Expressions:

Regular expressions are important components in developing PEX rules. Regular expressions or 'regex' are special text strings that describe patterns, these patterns require no linguistic analysis so they can be implemented easily across different language pairs. KantanMT have a list of regular expressions used for both GENTRY Rule files (*.rul) and PEX post-edit files (*.pex).